Claims

I claim:

1. A method for identifying moving objects in a video comprising:

segmenting the video into a plurality of volumes;

extracting a plurality of masks from the video; and

applying the plurality of masks to the plurality of volumes to identify volumes corresponding to moving objects.

2. The method of claim 1 wherein the video includes a plurality of frames and each frame includes a plurality of pixels, further comprising:

constructing a spatiotemporal data structure from the pixels of the frames, wherein each element in the data structure is a vector that includes color values of a pixel (x, y, t), where (x,y) are coordinates of the pixel in a particular frame t in the video;

selecting a vector with a minimum color gradient magnitude as a marker;

growing an unrefined around each marker;

removing the unrefined volume from the data structure; and

repeated the selecting, growing, and removing steps until the data structure is empty.

3. The method of claim 2 further comprising:

merging unrefined volumes that do not satisfy a minimum size requirement with a closest larger volume.

4. The method of claim 2 wherein the color gradient magnitude $|\nabla S|$ is determined by:

$$|\nabla S(x,y,t)| = |w_y(x^-,y,t) - w_y(x^+,y,t)| + |w_u(x,y^-,t) - w_u(y^+,t)| + |w_v(x,y,t^-) - w_v(x,y,t^+)|,$$

where $(\,)^-$ and $(\,)^+$ represent equal distances from a central pixel in a local neighborhood of pixels.

5. The method of claim 1, further comprising:

determining distances between a pixel in a window of a current frame and pixels in a window in an adjacent frame;

selecting a minimum distance of the distances as a score of each pixel in the window in adjacent frame;

averaging the scores; and

thresholding the scores to produce the masks.

6. The method of claim 1, further comprising:

using frame differences as change detection masks.

7. The method of claim 1, wherein the step of applying the plurality of masks to the plurality of volumes to identify moving objects further comprises:

counting the number of changed pixels in portions of the masks that intersect each volume; and

selecting volumes having counts exceeding a predetermined threshold as moving objects.

8. The method of claim 6, further comprising:

normalizing the total counts.